On the Detection of Imprinted Quantitative Trait Loci in Line Crosses: Effect of Linkage Disequilibrium

Cynthia Sandor and Michel Georges¹

Unit of Animal Genomics, Groupement Interdisciplinaire en Génomique Appliquée-Recherche and
Faculty of Veterinary Medicine, University of Liège, 4000-Liège, Belgium

Manuscript received June 11, 2008
Accepted for publication July 23, 2008

ABSTRACT

Imprinted quantitative trait loci (QTL) are commonly reported in studies using line-cross designs, especially in livestock species. It was previously shown that such parent-of-origin effects might result from the nonfixation of QTL alleles in one or both parental lines, rather than from genuine molecular parental imprinting. We herein demonstrate that if linkage disequilibrium exists between marker loci and nonfixed QTL, spurious detection of pseudo-imprinting is increased by an additional 40–80% in scenarios mimicking typical livestock situations. This is due to the fact that imprinting can be tested only in F₂ offspring whose sire and dam have distinct marker genotypes. In the case of linkage disequilibrium between markers and QTL, such parents have a higher chance to have distinct QTL genotypes as well, thus resulting in distinct padumal and madumal allele substitution effects, *i.e.*, QTL pseudo-imprinting.

A minority of mammalian genes (~100) are subject to parental imprinting (<http://www.har.mrc.ac.uk/mousebook/?by=imprinting>). Imprinted genes carry parent-of-origin-specific epigenetic marks that cause tissue- and developmental-stage-specific silencing of either the padumal (*i.e.*, transmitted by the father) (~50% of imprinted genes) or the madumal (*i.e.*, transmitted by the mother) (~50%) allele. Most imprinted genes cluster in imprinted domains, encompassing monoallelically (both padumal and madumal) as well as biallelically expressed genes. Monoallelically expressed genes at a given imprinted domain are coordinately regulated by imprinting control regions (ICR). In mammals, all known ICRs are characterized by allele-specific DNA methylation marks that in ~80% of the cases are imposed in the female germline (for instance, EDWARDS and FERGUSON-SMITH 2007). Note that recent bioinformatic analyses suggest that the number of imprinted genes may be higher but this remains to be demonstrated experimentally (LUEDI *et al.* 2005, 2007).

In animals, parental imprinting has been reported in placental mammals (marsupials and eutherians), but not in monotremes, birds, or fish (REIK and LEWIS 2005). Parental imprinting is thought to have evolved as a result of unequal parental provision of resources to offspring in polygamous species (the “conflict hypothesis”) (WILKINS and HAIG 2003). Padumal alleles benefit from levying maternal resources even if at the expense of future offspring of the mother, while madumal alleles benefit from sparing the reproductive

potential of the mother. Concomitantly, imprinted genes that are preferentially expressed from the padumal allele are enriched in genes that promote fetal and placental growth (*e.g.*, *Igf2* and *Peg11*), while imprinted genes that are preferentially expressed from the madumal allele are enriched in genes with opposite effects (*e.g.*, *Igf2r*, *Phlda2*, and *Cdkn1c*). There is evidence that some imprinted genes affect neonatal adaptation to suckling and metabolism (*Gnasxl* and *Peg3*) (CONSTANCIA *et al.* 2004), suggesting that they may be imprinted in monotremes as well, an hypothesis that remains to be tested.

It is noteworthy that parental imprinting has evolved independently in flowering plants, affecting genes expressed in the endosperm, an organ that provides nutrients to the embryo developing inside the seed (FEIL and BERGER 2007).

Growth and body composition are among the economically most important traits in livestock production, and identifying quantitative trait loci (QTL) and genes affecting these traits is one of the priorities of modern animal genetics (GEORGES 2007). Given the importance of imprinted genes in controlling prenatal growth, it is not surprising that animal geneticists have searched for parent-of-origin effects on QTL and genes influencing these traits. As a matter of fact, parent-of-origin effects were observed early on for QTL increasing muscle mass in Callipyge sheep (COCKETT *et al.* 1996) and Piétrain pigs (JEON *et al.* 1999; NEZER *et al.* 1999). The corresponding QTL were subsequently shown to be caused by postnatal overexpression of the imprinted *DLK1* (DAVIS *et al.* 2004) and *IGF2* genes (VAN LAERE *et al.* 2003) in skeletal muscle.

¹Corresponding author: Unit of Animal Genomics, GIGA-R Faculty of Veterinary Medicine, University of Liège, GIGA Tower (B34), 1 rue de l'Hôpital, 4000-Liège, Belgium. E-mail: michel.georges@ulg.ac.be

These initial studies were followed by a flurry of articles reporting imprinted QTL in domestic animals. As an example, DE KONING *et al.* (2000) detected four imprinted QTL (two paternally and two maternally expressed) of five affecting body composition in a Meishan $\times$ European F_2 and concluded that imprinting indeed plays an important role in the determinism of these traits. Reporting imprinted QTL has become standard in animal genetics, and the number of allegedly imprinted QTL abounds across the genome. Imprinted QTL have been reported even in poultry despite the fact that there is no experimental evidence for parental imprinting in birds (TUISKULA-HAAVISTO *et al.* 2004).

How can the present molecular knowledge of a handful of imprinted domains in mammals be reconciled with the abundance of imprinted QTL in domestic animals including birds? Studies reporting imprinted QTL share a “line-cross design,” *i.e.*, an F_2 intercross between lines or breeds diverging for the traits of interest. In these studies, parental lines/breeds are assumed to be fixed for alternate QTL alleles (Q and q), all F_1 individuals are assumed to be of “ Qq ” QTL genotype, and the three possible genotypes (QQ , Qq , qq) are assumed to segregate in the F_2 generation in $\sim 1:2:1$ Mendelian proportions. Testing imprinting consists of contrasting the madumal *vs.* padumal allele substitution effect at the QTL map position. Assuming that the marker alleles originating from the parental lines are labeled “1” and “2,” respectively, the madumal allele substitution effect is estimated as

$$\alpha_M = \frac{1}{2}[(\overline{11} - \overline{21}) + (\overline{12} - \overline{22})] \quad (1)$$

and the padumal allele substitution effect as

$$\alpha_P = \frac{1}{2}[(\overline{11} - \overline{12}) + (\overline{21} - \overline{22})]. \quad (2)$$

In these $\overline{XX}$ is the phenotypic average of offspring with corresponding marker genotype; the first allele is madumal, the second is padumal. The contrast that tests imprinting is

$$\Delta_{IMP} = \alpha_M - \alpha_P = (\overline{12} - \overline{21}). \quad (3)$$

As one cannot distinguish 12 from 21 offspring, imprinting cannot straightforwardly be tested in a typical F_2 setting—at least not using genotypic means.

Note that CUI *et al.* (2007) recently proposed to exploit known differences between male and female recombination rates to infer the most likely parental chromosomal origin in an F_2 generation. This information can in theory be used to test for imprinted QTL in an F_2 design. However, the effectiveness of this approach remains to be demonstrated, especially in species where differences between male and female recombination

rates are mostly subtle, including several livestock species.

How are animal geneticists then testing the imprinting status of QTL in line crosses? The previous conclusion assumes that alternate marker alleles (1 and 2) are fixed in the parental lines, as expected if these are inbred. However, lines or breeds of domestic animals are far from inbred, as demonstrated by the segregation of multiple marker alleles (*e.g.*, $1^1, 1^2, 1^3, \dots; 2^1, 2^2, 2^3, \dots$; note that there is a considerable degree of allele sharing between breeds/lines; *i.e.*, 1^m and 2^n may be identical-by-state for some m 's and n 's.). As a consequence F_1 parents often have distinct heterozygous marker genotypes. In these cases the two classes of heterozygous F_2 offspring ($1^m 2^n$ and $2^{n'} 1^{m'}$) can be distinguished and imprinting can be tested.

The caveat resides in the fact that animal geneticists typically assume that alternate QTL alleles are still fixed in the parental lines/breeds (even if marker alleles obviously are not) and that all F_1 individuals have Qq genotype (*e.g.*, HALEY *et al.* 1994). If this is not the case, F_1 animals may have different QTL genotypes, and the estimated QTL allele substitution effect will be the weighted average of allele substitution effects across F_1 parents. As the number of F_1 parents that is used to generate the F_2 population is typically limited—especially on the male side—the average madumal and padumal substitution effects may differ as a result of varying QTL genotype frequencies between F_1 males and females simply due to sampling. As recognized by DE KONING *et al.* (2002), there is thus a much more trivial explanation for the frequent parent-of-origin effects observed for QTL in domestic animals than genuine parental imprinting.

It has recently become apparent that linkage disequilibrium (LD) extends over much longer chromosome segments in breeds of domestic animals than in humans. Significant gametic associations between markers are readily detected using low-density microsatellite maps (*e.g.*, FARNIR *et al.* 2000; McRAE *et al.* 2002; NSENGIMANA *et al.* 2004; HARMEGNIES *et al.* 2006) and *a fortiori* using medium- or high-density SNP panels (*e.g.*, SUTTER *et al.* 2004; JUNGIERUS *et al.* 2005; LINDBLAD-TOH *et al.* 2005). We reasoned that if LD occurs between markers and QTL—as expected—the imprinting hypothesis will preferentially be testable in offspring from matings between F_1 parents that differ for their QTL genotype as a result of their distinct marker genotype. This would irretrievably result in a difference between the madumal and the padumal allele substitution effect within families, *i.e.*, a QTL with parent-of-origin effect. But would this hypothesis translate in an increase in “spurious” QTL imprinting across families? To address this question, we herein examine the effect of LD on the incidence of parent-of-origin effects when mapping QTL in line-cross designs typical of studies in domestic animals.

METHODS

Simulations: We first produced a “base” population by randomly mating 1000 males and 1000 females for 10,000 generations using a recombination–mutation–drift model. Twenty-one unlinked QTL were “dropped” in this population, with a mutation rate $\mu_{\text{QTL}} = 10^{-5}$ under an approximate infinite-alleles model (maximum number of alleles = 1000). Each QTL was positioned at position 22 cM of a 50-cM chromosome (a different chromosome for each QTL) with six evenly spaced markers characterized by a mutation rate $\mu_{\text{M}} = 10^{-3}$ under the same approximate infinite-alleles model. QTL and marker were assumed to be monomorphic in the first generation.

At generation 10,000, we assigned an allelic effect to each QTL allele in the population. QTL effects were sampled in a double-negative exponential distribution, $e^{-\delta x}/2\delta$, and randomly assigned a positive or a negative sign (SANDOR *et al.* 2006). The shape parameter, δ , was set at one.

The “breeding value” (BV) of each individual was computed as the sum of effects across all its constituent—padumal and madumal—QTL alleles. The simulated QTL were thus considered to act in a Mendelian fashion; *i.e.*, none of them were imprinted. In addition, each individual was assigned a nongenetic effect sampled from a normal distribution $N(0, \sigma_{\text{E}}^2)$. The value of σ_{E}^2 was set such that $h^2 = \sigma_{\text{BV}}^2 / (\sigma_{\text{BV}}^2 + \sigma_{\text{E}}^2)$ was either 0.4 or 0.8 at generation 10,000. It remained unchanged in subsequent generations (see hereafter). The individual’s nongenetic effect was added to the breeding value to yield its “phenotypic value,” y_i .

From this base population we then generated two distinct populations (each with 200 males and 500 females) by divergent selection on the phenotypic values. This was achieved by assigning differential probabilities of parenthood to individuals as a function of their phenotype. Probabilities of parenthood of individual i of sex s were computed as

$$P_{is,H} = \frac{(y_{is} - \text{Min} + 1)^\lambda}{\sum_{i=1}^n (y_{is} - \text{Min} + 1)^\lambda} \quad (4)$$

and

$$P_{is,L} = \frac{(\text{Max} - y_{is} + 1)^\lambda}{\sum_{i=1}^n (\text{Max} - y_{is} + 1)^\lambda} \quad (5)$$

in the population selected for “high” ($P_{is,H}$) *vs.* “low” ($P_{is,L}$) phenotypic value, respectively. Min and Max are the lowest and highest phenotypic values in the respective populations, and λ determines the selection intensity. Selection was applied for up to 20 generations. Mutation and recombination operated during selection as before.

From these selected populations, we sampled 20 F_0 sires in one line and 80 F_0 dams in the other. The choice of lines (high or low) from which to sample the F_0 sires

or dams was random. Each F_0 sire was then mated to 4 F_0 dams, producing five offspring each, for a total of 400 F_1 . From these we randomly sampled 80 F_1 dams to generate five F_2 offspring each by equitably mating them with (i) 2 F_1 sires, (ii) 10 F_1 sires, or (iii) 20 F_1 sires, for a total of 400 F_2 offspring. This procedure was in essence identical to the one used by DE KONING *et al.* (2000).

For each simulation, we selected the QTL contributing most to the genetic variance, σ_{BV}^2 , for linkage disequilibrium analysis and QTL mapping (see hereafter). The selected QTL are referred to as the “target QTL” in the remainder of the article.

Linkage disequilibrium: The simulation procedure generates LD between markers and between the markers and the QTL. Pairwise LD was quantified using r^2 as described (GRISART *et al.* 2004). To neutralize the effect of LD and hence be able to rigorously quantify its effect on the detection of “imprinted” QTL, we randomly permuted marker genotypes within F_0 sires and dams, respectively, prior to “dropping” the chromosomes down the F_1 and F_2 generations. The latter comprised exactly the same individuals as before, thus not altering the distribution of QTL genotypes in the pedigree.

QTL mapping: QTL mapping of the target QTL was performed by least-squares analysis, using the line-cross model proposed HALEY *et al.* (1994) and KNOTT *et al.* (1998). Briefly, the presence of a Mendelian QTL at a given map position, j , was tested assuming that

$$y_i = \mu + a(p_{11}^{ij} - p_{22}^{ij}) + d(p_{12}^{ij} + p_{21}^{ij}) + \varepsilon_{ij} \quad (6)$$

while the presence of an imprinted QTL was tested assuming that

$$y_i = \mu + a(p_{11}^{ij} - p_{22}^{ij}) + d(p_{12}^{ij} + p_{21}^{ij}) + m(p_{12}^{ij} - p_{21}^{ij}) + \varepsilon_{ij}. \quad (7)$$

In these y_i is the phenotypic value of individual i , μ is the midpoint between the phenotypic value of alternate homozygotes, p_{xy}^{ij} is the probability that individual i has genotype “ xy ” at map position j (where x is the madumal allele and y the padumal allele) conditional on flanking marker genotypes, a is the additive effect, d the dominance deviation, m the imprinting deviation, and ε_{ij} the error term for individual i at map position j . The values of μ , a , d , and m are determined to minimize the sum of squared error (SSE) terms over all offspring. Under the null hypothesis of no QTL

$$\frac{\text{SSR}/k}{\text{SSE}/(n - k - 1)} \quad (8)$$

has an F -distribution with k and $n - k - 1$ d.f. SSR corresponds to the sum of squares due to the model, n to the number of offspring, and k to the number of parameters to estimate, *i.e.*, two in the Mendelian scenario and three when testing imprinting.

FIGURE 1.—(A) Evolution of the proportion of polymorphic loci (major allele frequency <0.95) (circles) and rate of heterozygosity (squares) for marker loci (red) and QTL (black) during the creation of the base population (10,000 generations). Data points are shown every 100 generations. Each data point corresponds to the average over 6 (markers), respectively 21 (QTL) loci and 10,000 simulations. (B) Evolution of the same metrics in the selection lines (20 generations) in the presence ($\lambda = 2$) and the absence ($\lambda = 0$; miniatures) of selection. The data for the target QTL are shown in blue. Every data point corresponds to the average across 6 (markers), 1 (target QTL), and 20 (background QTL) loci times 10,000 simulations. (C) Evolution of LD [$\log_{10}(r^2)$] between markers and target QTL during creation of the base population. Data points correspond to the average of 1000 simulations and are—for clarity—shown only for the marker at 2 cM (red circles) and the marker at 18 cM (black circles) from the QTL. (D) Same LD metrics in the selection lines (20 generations), with ($\lambda = 2$) and without ($\lambda = 0$; miniatures) selection. The effect of genotype permutations on LD is shown for the same two-marker-QTL combinations (red and black squares). Every data point corresponds to the average across 1000 simulations.

Nominal P -values of the F -statistics were corrected for multiple testing (51 positions and two models) by permutation testing (CHURCHILL and DOERGE 1994). The corrected P -value corresponds to their rank (divided by 1000) with respect to the lowest P -values (across positions and models) obtained across 1000 permuted data sets. The threshold for significance was set at 0.05.

For data sets yielding significant QTL, if the most significant P -value on the chromosome was obtained with the Mendelian model, the QTL was assumed to be Mendelian. If the most significant P -value was obtained with the imprinting model, we rejected the Mendelian model in favor of the imprinting model only if at the most significant position of the imprinted QTL

$$\frac{\text{SSR}_{\text{IMP}} - \text{SSR}_{\text{MEND}}}{\text{SSE}_{\text{MEND}}/(n-4)} = F_{1,n-4} \quad (9)$$

had a P -value <0.05 .

RESULTS

Figure 1A shows the progression of the level of polymorphism during the creation of the base population ($N_e = 2000$). The proportion of polymorphic loci, defined as loci with major allele frequency <0.95 , increases to 0.90 for marker loci and 0.69 for QTL. Concomitantly the rate of heterozygosity plateaus at 0.44 and 0.28 for markers and QTL, respectively, complying with theoretical expectation: $H_e = 4N_e\mu/(4N_e\mu + 1)$.

Figure 1B shows the evolution of the same metrics in the selection lines ($N_e = 572$). Without selection ($\lambda = 0$), the proportion of polymorphic loci and the rate of heterozygosity decrease slowly toward equilibrium values dictated by effective population size. With selection ($\lambda = 2$), the reduction of polymorphism is accelerated at all QTL and particularly at the target QTL selected on the basis of its largest contribution to genetic variance. The proportion of populations for which the target

FIGURE 2.—(A) Frequency distribution (over 100 simulations) of QTL allelic effects (absolute value) for the target QTL (red), as well as across the 20 remaining background QTL (black). (B) Frequency distribution of the proportion of genetic variance explained by the target QTL (red) as well as by individual background QTL.

QTL is polymorphic decreases from ~ 0.70 to ~ 0.35 in 20 generations. Under selection, the decrease of polymorphism is also accelerated at marker loci, particularly those that undergo the strongest hitchhiking effects due to their proximity to the target QTL (data not shown).

Figure 1C shows the evolution of LD between the target QTL and linked markers in the base population. As expected, LD stabilizes at equilibrium values in accordance with theory: $r^2 = 1/(4N_e\theta + 1)$. On creation of the selection lines, r^2 values first undergo a slight increase due to reduction in population size ($+1/n$). In the absence of selection ($\lambda = 0$), LD then increases toward new equilibrium values, while selection ($\lambda = 2$) clearly augments LD buildup. Figure 1D also shows the effect of marker genotype permutations: LD becomes independent of genetic distance and is determined exclusively by population size.

Figure 2A shows the frequency distribution of allelic effects for target QTL as well as for the remaining 20 “background” QTL. Target QTL are typically characterized by one allele with a large substitution effect and

an average of 3.4 alleles with a modest effect (supplemental Figure 1). This situation is reminiscent of what has been observed in reality at, for instance, the *DGAT1*, *GHR*, and *ABCG2* loci in cattle (for instance, GRISART *et al.* 2002, 2004; BLOTT *et al.* 2003; COHEN-ZINDER *et al.* 2005), the *MSTN* locus in sheep (CLOP *et al.* 2006), and the *IGF2* locus in pigs (VAN LAERE *et al.* 2003). Many QTL are indeed likely to be polyallelic *sensu stricto*, yet likely to behave largely as biallelic because of the occurrence of one allele with an unusually large allelic effect. Figure 2B shows the frequency distribution of the proportion of the genetic variance accounted for by the target QTL as well as of the remaining background QTL. The average proportion of genetic variance explained was 31% for the target QTL and 2% for a typical background QTL.

Figure 3, A–C, shows the evolution of the power to detect the target QTL—using a line-cross model—as a function of (i) the number of generations of selection (1–20), (ii) the heritability of the trait (0.40 or 0.80), and (iii) the F_2 pedigree structure ($2\delta/80\varnothing$, $10\delta/80\varnothing$, and $20\delta/80\varnothing$). As expected, detection power is very low at the onset of selection as both populations have not yet diverged and the average substitution effects between QTL alleles originating from the “low” and “high” line are concomitantly close to zero in the F_2 population. Nevertheless, at selection onset power is slightly higher under the $2\delta/80\varnothing$ scenario, supposedly reflecting situations where at least one sire is or both are by chance “concordantly” heterozygous at the QTL. Detection power systematically increases with the number of generations of selection under all scenarios, as a result of the increasing divergence of the parental lines at the target QTL. It reaches slightly higher levels under the $10\delta/80\varnothing$ and $20\delta/80\varnothing$ scenarios than under the $2\delta/80\varnothing$ scenario. The reason for this remains uncertain but may be due at least in part to a higher variance in marker information content across simulations in the $2\delta/80\varnothing$ than in the other scenarios. As expected, decreased heritability decreases detection power under all scenarios as the additional nongenetic noise reduces the ratio of QTL variance to phenotypic variance and hence the ability to detect the QTL.

Strikingly, the proportion of cases in which one erroneously concludes that a detected QTL (*i.e.*, exceeding the 0.05 significance threshold as defined in METHODS) is imprinted is $>60\%$ (average: 84%) for the first 10 generations of selection under the $2\delta/80\varnothing$ scenario, $>20\%$ (average: 34%) under the $10\delta/80\varnothing$ scenario, and $>15\%$ (average: 24%) under the $20\delta/80\varnothing$ scenario (Figure 3, A–C). The decrease in the proportion of erroneous conclusions of imprinting with increasing numbers of F_1 sires agrees with the findings of DE KONING *et al.* (2002), reflecting a decrease in the standard error of the estimate of the average paternal QTL allele substitution effect with increasing number of F_1 sires. Erroneous detection of imprinting decreases with increasing generations of selection, yet remains

FIGURE 3.—(A–C) Power to detect target QTL (triangles), and incidence of erroneous conclusions of imprinting among detected QTL (circles), as a function of the number of generations of selection (1–20), in the presence (red) and absence (black) of LD between markers and QTL, under the 2♂/80♀ (A), 10♂/80♀ (B), and 20♂/80♀ (C) F₁ scenarios, for a trait with heritability of 40% (solid symbols and solid lines) or 80% (open symbols and dotted lines). Data points correspond to averages over 10,000 simulations. (D–F) Proportion of simulations for which the average QTL allele substitution effects *within* lines are superior to the average QTL allele substitution effects between lines (triangles). The incidence of erroneous conclusions of imprinting among detected QTL (circles) in the presence (red) and absence (black) of LD between markers and QTL is shown, under the 2♂/80♀ (A), 10♂/80♀ (B), and 20♂/80♀ (C) F₁ scenarios, for a trait with heritability of 40% (solid symbols and solid lines) or 80% (open symbols and dotted lines). Data points correspond to averages over 10,000 simulations.

>18, 9, and 8% at generation 20 under the 2♂/80♀, 10♂/80♀, and 20♂/80♀ scenarios, respectively. This decrease is presumably due to the fact that alternate QTL alleles are getting near fixation in the high and low lines in an increasing proportion of simulations, thereby reducing the variation in QTL genotype among F₁ parents—the source of the problem. Indeed, when restricting the analyses to crosses between selected lines for which the average QTL substitution effect *within* lines is larger than the average substitution effect *between* lines, the proportion of erroneous imprinting calls remains very high: 88% on average under the 2♂/80♀

scenario, 36% on average under the 10♂/80♀ scenario, and 26% on average under the 20♂/80♀ scenario (Figure 3, D–F).

Erroneous detection of imprinting increased slightly with decreasing heritability (Figure 3, A–C, and data not shown). We believe that this reflects less effective selection and hence more prolonged segregation of distinct QTL alleles in the high and low lines. Indeed, this effect of heritability disappeared when restricting the analyses to crosses between selected lines for which the average QTL substitution effect *within* lines is larger than the average substitution effect *between* lines (Figure 3, D–F).

FIGURE 3.—Continued.

The contribution of LD to erroneous detection of imprinting—the main hypothesis tested in this study—is clearly demonstrated by examining the effect of marker genotype permutation to neutralize LD between markers and QTL in the selection lines prior to mating (Figure 3). While erroneous imprinting calls indeed occur in the absence of LD as well, corroborating the results of DE KONING *et al.* (2002), with LD the incidence of this artifact is increased by $\sim 40\%$ on average under the 2♂/80♀ scenario, by $\sim 70\%$ under the 10♂/80♀ scenario, and by $\sim 80\%$ under the 20♂/80♀ scenario. These results thus clearly show that—at least under the studied scenarios—LD considerably exacerbates the problem, as surmised. Note that the QTL detection power is unaffected by the presence or the absence of LD (Figure 3).

DISCUSSION

QTL mapping in domestic animals is routinely performed in F_2 populations derived from parental lines that are divergent for at least some of the analyzed traits. The most commonly applied statistical models assume that the parental lines are fixed for alternate QTL alleles. Yet it is becoming increasingly apparent that in many cases this assumption is not valid: parental lines are often segregating for the QTL as well as sharing QTL alleles. With hindsight and having identified the causal mutation, this is now clearly demonstrated for the *IGF2*-intron-nt3072 QTN (VAN LAERE *et al.* 2003).

Applying a line-cross model to such data is suboptimal and may lead to loss of QTL detection power. Moreover, as previously noticed by DE KONING *et al.* (2002) and extended in this work, it may lead to erroneous conclusions about the mode of action of the QTL, particularly with regard to parental imprinting. DE KONING *et al.* (2002) realized that if the parental lines are not fixed for alternate QTL alleles, the “average” QTL genotype of the utilized F_1 sires and dams may differ as a result of sampling, thus resulting in a difference between the padumal and madumal QTL allele substitution effects even in the absence of true parental imprinting. We herein demonstrate that if LD exists between the markers and the QTL, “pseudo-imprinting” may become the rule rather than the exception in scenarios that are likely to properly mimic real livestock situations. We surmise that this is due to the fact that in the presence of LD the padumal and madumal QTL allele substitution effects are more effectively contrasted in matings between parents with distinct marker and hence QTL genotype. We believe that this accounts for the fact that imprinted QTL effects have been so commonly observed in line crosses.

It may be argued that many commercial populations have been selected for >20 generations, the time horizon considered in our simulations. However, selection objectives often change over time, adjusting to evolving economic constraints. As an example, it is likely that the *DGATI* K232 allele was initially selected for in dairy cattle because of its favorable effect on milk yield, until fat yield became the primary objective, hence favoring the alternate 232A allele, while this polymorphism is now neutral with respect to the present-day selection Dutch index that weights both fat and protein yield, thus explaining its continued segregation (GRISART *et al.* 2002). Moreover, QTL mapping experiments often target traits that are of major scientific interest without necessarily being the focus of ongoing breeding programs. Thus segregation of QTL alleles is likely to be a common occurrence in parental livestock populations used to generate line crosses.

Such crosses are thus better analyzed using models that do not make the assumption that all alleles originating from a given parental line have the same QTL

effect. This is sometimes achieved by analyzing line-cross data using a half-sib design, taking advantage of the common occurrence of large paternal half-sib pedigrees. This approach, however, foregoes the information from the maternal chromosomes. More appropriate is the use of variance component (VC) approaches that either exploit linkage information alone (for instance, GEORGE *et al.* 2000) or simultaneously exploit linkage and LD signal to compute the probability of identity-by-descent (IBD) between chromosomes, used to constrain the covariances between the corresponding haplotype effects (for instance, ULEBERG *et al.* 2005).

Parent-of-origin effects could be tested within a VC framework by estimating the variance associated with the madumal and padumal haplotypes separately (for instance, HEUVEN *et al.* 2005) and comparing the likelihood of the data assuming that these are different *vs.* identical. However, this approach would suffer from the same drawback as the line-cross model in that the identification of a statistically different madumal *vs.* padumal VC does not imply that genuine parental imprinting is involved.

Demonstrating genuine imprinting thus requires the comparison of the QTL allele substitution effect of a proven IBD pair of alleles upon paternal *vs.* maternal transmission. This is difficult to achieve in livestock as it is only exceptionally possible to have a sufficiently large number of F₁ dams that have a genotype that is unambiguously IBD with that of one or more F₁ sires.

The availability of hundreds of thousands to millions of SNPs for most livestock species, combined with cost-effective high-throughput screening techniques, will soon allow the recognition of haplotypes that are known with virtual certainty to be IBD even in the absence of pedigree data. It should thus become possible to test the imprinting hypothesis by comparing the phenotypes of alternate heterozygotes ("12's *vs.* 21's") at the population level. However, and as noted by others, this approach hides potential pitfalls as well, as maternal effects may render this contrast significant in the absence of genuine parental imprinting (HAGER *et al.* 2008).

Imprinted genes certainly contribute to the genetic variation of quantitative traits (*e.g.*, COCKETT *et al.* 1996; VAN LAERE *et al.* 2003), and forward genetics, including QTL mapping, may contribute to the identification of novel imprinted genes (*e.g.*, WOLF *et al.* 2008). However, it is important to recognize that alternative "artifactual" explanations may account for the parent-of-origin effects that are often found for QTL detected in line crosses. Demonstrating a role for genuine parental imprinting requires more stringent tests than what is usually applied. For certain, available QTL evidence does not warrant reevaluation of how common parental imprinting is both in terms of the number of affected genes and the species in which it occurs.

Cynthia Sandor is a fellow of the "Fonds pour la Recherche dans l'Industrie et l'Agriculture." Support for this work was provided by the Walloon Ministry of Agriculture, the Belgian Science Policy organiza-

tion (Biomagnet Pôle d'Attraction Interuniversitaire), and the Communauté Française de Belgique (Biomod Action de Recherche Concertée).

LITERATURE CITED

- BLOTT, S., J. J. KIM, S. MOISIO, A. SCHMIDT-KÜNTZEL, A. CORNET *et al.*, 2003 Molecular dissection of a quantitative trait locus: a phenylalanine-to-tyrosine substitution in the transmembrane domain of the bovine growth hormone receptor is associated with a major effect on milk yield and composition. *Genetics* **163**: 253–266.
- CLOP, A., F. MARCQ, H. TAKEDA, D. PIROTTIN, X. TORDOIR *et al.*, 2006 A mutation creating a potential illegitimate microRNA target site in the myostatin gene affects muscularity in sheep. *Nat. Genet.* **38**: 813–818.
- CHURCHILL, G. A., and R. W. DOERGE, 1994 Empirical threshold values for quantitative trait mapping. *Genetics* **138**: 963–971.
- COCKETT, N. E., S. P. JACKSON, T. L. SHAY, F. FARNIR, S. BERGHMANS *et al.*, 1996 Polar overdominance at the ovine callipyge locus. *Science* **273**: 236–238.
- COHEN-ZINDER, M., E. SEROUSSI, D. M. LARKIN, J. J. LOOR, A. EVERTS-VAN DER WIND *et al.*, 2005 Identification of a missense mutation in the bovine ABCG2 gene with a major effect on the QTL on chromosome 6 affecting milk yield and composition in Holstein cattle. *Genome Res.* **15**: 936–944.
- CONSTANCIA, M., G. KELSEY and W. REIK, 2004 Resourceful imprinting. *Nature* **432**: 53–57.
- CUI, Y., J. M. CHEVERUD and R. WU, 2007 A statistical model for dissecting genomic imprinting through genetic mapping. *Genetica* **130**: 227–239.
- DAVIS, E., C. H. JENSEN, H. D. SCHRODER, F. FARNIR, T. SHAY-HADFIELD *et al.*, 2004 Ectopic expression of DLK1 protein in skeletal muscle of padumal heterozygotes causes the callipyge phenotype. *Curr. Biol.* **14**: 1858–1862.
- DE KONING, D. J., A. P. RATTINK, B. HARLIZIUS, J. A. VAN ARENDONK, E. W. BRASCAMP *et al.*, 2000 Genome-wide scan for body composition in pigs reveals important role of imprinting. *Proc. Natl. Acad. Sci. USA* **97**: 7947–7950.
- DE KONING, D. J., H. BOVENHUIS and J. A. VAN ARENDONK, 2002 On the detection of imprinted quantitative trait loci in experimental crosses of outbred species. *Genetics* **161**: 931–938.
- EDWARDS, C. A., and A. C. FERGUSON-SMITH, 2007 Mechanisms regulating imprinted genes in clusters. *Curr. Opin. Cell Biol.* **19**: 281–289.
- FARNIR, F., W. COPPIETERS, J. J. ARRANZ, P. BERZI, N. CAMBISANO *et al.*, 2000 Extensive genome-wide linkage disequilibrium in cattle. *Genome Res.* **10**: 220–227.
- FEIL, R., and F. BERGER, 2007 Convergent evolution of genomic imprinting in plants and mammals. *Trends Genet.* **23**: 192–199.
- GEORGE, A. W., P. M. VISSCHER and C. S. HALEY, 2000 Mapping quantitative trait loci in complex pedigrees: a two-step variance component approach. *Genetics* **156**: 2081–2092.
- GEORGES, M., 2007 Mapping, fine-mapping and molecular dissection of quantitative trait loci in domestic animals. *Annu. Rev. Genomics Hum. Genet.* **8**: 131–162.
- GRISART, B., W. COPPIETERS, F. FARNIR, L. KARIM, C. FORD *et al.*, 2002 Positional candidate cloning of a QTL in dairy cattle: identification of a missense mutation in the bovine DGAT gene with major effect on milk yield and composition. *Genome Res.* **12**: 222–231.
- GRISART, B., F. FARNIR, L. KARIM, N. CAMBISANO, J. J. KIM *et al.*, 2004 Genetic and functional confirmation of the causality of the DGAT1 K232A quantitative trait nucleotide in affecting milk yield and composition. *Proc. Natl. Acad. Sci. USA* **101**: 2398–2403.
- HAGER, R., J. M. CHEVERUD, C. ROSEMAN and J. B. WOLF, 2008 Maternal effects as the cause of parent-of-origin effects that mimic genomic imprinting. *Genetics* **178**: 1755–1762.
- HALEY, C. S., S. A. KNOTT and J. M. ELSN, 1994 Mapping quantitative trait loci in crosses between outbred lines using least squares. *Genetics* **136**: 1195–1207.
- HARMEGNIES, N., F. FARNIR, F. DAVIN, N. BUYS, M. GEORGES *et al.*, 2006 Measuring the extent of linkage disequilibrium in commercial pig populations. *Anim. Genet.* **37**: 225–231.
- HEUVEN, H.C.M., H. BOVENHUIS, L. L. G. JANSSEN and J. A. M. VAN ARENDONK, 2005 Efficiency of population structures for mapping of Mendelian and imprinted quantitative trait loci in outbred

- pigs using variance component methods. *Genet. Sel. Evol.* **37**: 635–655.
- JEON, J. T., O. CARLBORG, A. TORNSTEN, E. GIUFFRÀ, V. AMARGER *et al.*, 1999 A paternally expressed QTL affecting skeletal and cardiac muscle mass in pigs maps to the IGF2 locus. *Nat. Genet.* **21**: 157–158.
- JUNGERIUS, B. J., J. GU, R. P. CROOIJMANS, J. J. VAN DER POEL, M. A. GROENEN *et al.*, 2005 Estimation of the extent of linkage disequilibrium in seven regions of the porcine genome. *Anim. Biotechnol.* **16**: 41–54.
- KNOTT, S. A., L. MARKLUND, C. S. HALEY, K. ANDERSSON, W. DAVIES *et al.*, 1998 Multiple marker mapping of quantitative trait loci in a cross between outbred wild boar and large white pigs. *Genetics* **149**: 1069–1080.
- LINDBLAD-TOH, K., C. M. WADE, T. S. MIKKELSEN, E. K. KARLSSON, D. B. JAFFE *et al.*, 2005 Genome sequence, comparative analysis and haplotype structure of the domestic dog. *Nature* **438**: 803–819.
- LUEDI, P. P., A. J. HARTEMINK and R. L. JIRTLE, 2005 Genome-wide prediction of imprinted murine genes. *Genome Res.* **15**: 875–884.
- LUEDI, P. P., F. S. DIETRICH, J. R. WEIDMAN, J. M. BOSKO, R. L. JIRTLE *et al.*, 2007 Computational and experimental identification of novel human imprinted genes. *Genome Res.* **17**: 1723–1730.
- MCRÆ, A. F., J. C. MCEWAN, K. G. DODDS, T. WILSON, A. M. CRAWFORD *et al.*, 2002 Linkage disequilibrium in domestic sheep. *Genetics* **160**: 1113–1122.
- NEZER, C., L. MOREAU, B. BROUWERS, W. COPPIETERS, J. DETILLEUX *et al.*, 1999 An imprinted QTL with major effect on muscle mass and fat deposition maps to the IGF2 locus in pigs. *Nat. Genet.* **21**: 155–156.
- NSENGIMANA, J., P. BARET, C. S. HALEY and P. M. VISSCHER, 2004 Linkage disequilibrium in the domesticated pig. *Genetics* **166**: 1395–1404.
- REIK, W., and A. LEWIS, 2005 Co-evolution of X-chromosome inactivation and imprinting in mammals. *Nat. Rev. Genet.* **6**: 403–410.
- SANDOR, C., F. FARNIR, S. HANSOUL, W. COPPIETERS, T. MEUWISSEN *et al.*, 2006 Linkage disequilibrium on the bovine X chromosome: characterization and use in quantitative trait locus mapping. *Genetics* **173**: 1777–1786.
- SUTTER, N. B., M. A. EBERLE, H. G. PARKER, B. J. PULLAR, E. F. KIRKNESS *et al.*, 2004 Extensive and breed-specific linkage disequilibrium in *Canis familiaris*. *Genome Res.* **14**: 2388–2396.
- TUISKULA-HAAVISTO, M., D. J. DE KONING, M. HONKATUKIA, N. F. SCHULMAN, A. MAKI-TANILA *et al.*, 2004 Quantitative trait loci with parent-of-origin effects in chicken. *Genet. Res.* **84**: 57–66.
- ULEBERG, E., I. S. WIDERØE, E. GRINDFLEK, J. SZYDA, S. LIEN *et al.*, 2005 Fine mapping of a QTL for intramuscular fat on porcine chromosome 6 using combined linkage and linkage disequilibrium mapping. *J. Anim. Breed. Genet.* **122**: 1–6.
- VAN LAERE, A. S., M. NGUYEN, M. BRAUNSCHWEIG, C. NEZER, C. COLLETTE *et al.*, 2003 A regulatory mutation in IGF2 causes a major QTL effect on muscle growth in the pig. *Nature* **425**: 832–836.
- WILKINS, J. F., and D. HAIG, 2003 What good is genomic imprinting: the function of parent-specific gene expression. *Nat. Rev. Genet.* **4**: 359–368.
- WOLF, J. B., J. M. CHEVERUD, C. ROSEMAN and R. HAGER, 2008 Genome-wide analysis reveals a complex pattern of genomic imprinting in mice. *PLoS Genet.* **4**: e1000091.

Communicating editor: C. HALEY